# Solution Structure Determination and Mutational Analysis of the Papillomavirus E6 Interacting Peptide of E6AP[†,‡]

Xiaobing Be,[§] Yihui Hong,[||] Julia Wei,[§] Elliot J. Androphy,[||] Jason J. Chen,[||] and James D. Baleja*[,§]

*Department of Biochemistry, Tufts University School of Medicine, 136 Harrison Avenue, Boston, Massachusetts 02111, and Department of Dermatology, Box 166, New England Medical Center, 750 Washington Street, Boston, Massachusetts 02111*

*Received August 17, 2000; Revised Manuscript Received October 24, 2000*

ABSTRACT: E6AP is a cellular protein that binds cancer-related papillomaviral E6 proteins. The E6 binding domain, called E6ap, is located on an 18-amino acid segment of E6AP. The corresponding peptide was synthesized and its structure determined by nuclear magnetic resonance spectroscopy. The overall structure of the peptide is helical. A consensus E6-binding sequence among different E6 interacting proteins contains three conserved hydrophobic residues. In the structure of the E6AP peptide, the three conserved leucines (Leu 9, Leu 12, and Leu 13) form a hydrophobic patch on one face of the α-helix. Substitution of any of these leucines with alanine abolished binding to E6 protein, indicating that the entire hydrophobic patch is necessary. Mutation of a glutamate to proline, but not alanine, also disrupted the interaction between E6 and E6AP protein, suggesting that the E6-binding motif of the E6AP protein must be helical when bound to E6. Comparison of the E6ap structure and mutational results with those of another E6-binding protein (E6BP/ERC-55) indicates the existence of a general E6-binding motif.

The human papillomavirus (HPV)[1] is a small DNA tumor virus that infects epithelial cells (*1*). More than ninety different HPV types have been described. HPVs that infect the genital tract are divided into two groups (low-risk and high-risk) depending on the tendency of the type to progress to cancer. HPV type 16 (HPV-16), for example, is one strain in the high-risk group that is strongly associated with cervical carcinoma (*2*). Infection by a low-risk HPV strain can cause the cellular proliferation characteristic of warts but rarely leads to cancer (*3*). Viral infection occurs in the replicating basal layer epithelial cells, whereas viral replication is restricted to the upper epithelial layers where cells are committed to differentiation. The viral genome encodes a set of viral proteins that alter the differentiation pathway of the host cells for viral production. The viral E6 and E7 proteins of the high-risk papillomaviruses, which have high oncogenic potential, have identified roles in inhibiting differentiation and activating cellular proliferation. E7 inhibits the functions of the cell cycle repressor proteins P21 and RB, whereas E6 regulates protein degradation of the cellular tumor suppressor P53. By inhibiting cell cycle repressors, E6 and E7 function to promote viral DNA replication and viral protein synthesis (*2*).

A 100 kDa cellular protein, E6AP, is required for E6-dependent p53 degradation. E6AP, which contains ubiquitin-protein ligase activity in its C-terminal Hect domain, forms a stable complex with E6 that results in the specific ubiquitination and subsequent degradation of P53 (*4*). The E6 binding domain of E6AP has been localized to an internal segment comprising amino acid residues 391−408. This 18-amino acid segment, called E6ap, is both necessary and sufficient for E6 binding (*5*). E6AP is a cytoplasmic protein in which certain mutations (mostly in the Hect domain) are associated with Angelman's syndrome (*6*). Because E6 is not an abundant protein, its cellular location has been difficult to define, but it has been located in the cytoplasm and nucleus or associated with membranous compartments such as the ER (*3*).

Besides E6AP, other E6 interacting proteins have been identified that have a variety of functions. E6BP (ERC-55) is a calcium-binding protein located in the endoplasmic reticulum (*7*). Tuberin is a putative tumor suppressor protein (*8*). Paxillin associates with focal adhesion kinase and is implicated in cytoskeleton regulation (*9, 10*). IRF3 (interferon-regulatory factor 3) is a transcriptional factor (*11*). Sequence comparison among these E6 interacting proteins reveals a consensus sequence, L̲hxϕL̲sh (Figure 1), where the shaded letters indicate the most conserved and hydrophobic residues, h is an amino acid residue capable of accepting hydrogen bonds (D, E, Q, or N), x denotes any amino acid, ϕ is hydrophobic, and s represents a small amino acid residue (G or A). In a mutagenesis study performed on E6BP, substitution of any hydrophobic residue in the consensus sequence with alanine entirely abolished the binding between

[†] Supported by the Mary Kay Ash charitable foundation and NIH Grants AI38001 and R0173558. J.J.C. was supported in part by grants from the Dermatology Foundation Dermik Laboratories and the Cancer Research Foundation of America.

[‡] The atomic coordinates have been deposited with the Research Collaboratory for Structural Bioinformatics Protein Databank (entry 1EQX).

* To whom correspondence should be addressed. Phone: (617) 636-6872. Fax: (617) 636-2409. E-mail: jbaleja@opal.tufts.edu.

[§] Tufts University School of Medicine.

[||] New England Medical Center.

[1] Abbreviations: PV, papillomavirus; BPV, bovine papillomavirus; HPV, human papillomavirus; GST, glutathione *S*-transferase; TFE, trifluoroethanol; LSAB, low-salt association buffer; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; TOCSY, total correlation spectroscopy; rms, root-mean-square; CD, circular dichroism.

10.1021/bi0019592 CCC: $20.00 
Published on Web 01/09/2001

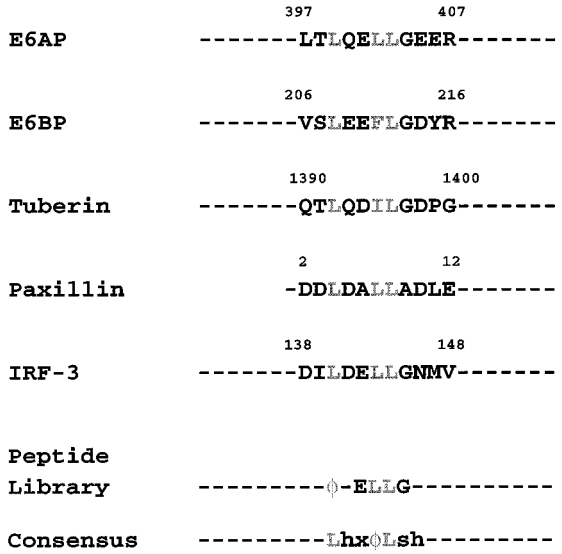


FIGURE 1: Schematic representation of E6-binding sequences. The hydrophobic amino acid residues conserved for E6 binding are shown in gray. The numbers on top indicate the sequence positions of these amino acid residues in the corresponding proteins.

BPV-1 E6 and E6BP, whereas replacement of the other consensus amino acid residues had a modest effect (*12*).

The three-dimensional structure of the E6 binding domain of the E6AP protein may provide insight into the interactions between E6 and its interacting proteins. In this paper, we present the solution structure of the E6ap peptide determined by nuclear magnetic resonance (NMR) methods. On the basis of this structure, a mutational study was performed to determine the importance of individual amino acid residues of E6AP for binding to the E6 of bovine papillomavirus (BPV-1) and human papillomavirus type 16 (HPV-16).

## EXPERIMENTAL PROCEDURES

*Peptide Synthesis.* An 18-amino acid residue peptide, E6ap, was prepared synthetically at the Tufts Peptide Facility using Fmoc [*N*-(9-fluorenyl)methoxycarbonyl] chemistry. The N-terminus was acetylated and the C-terminus amidated (*13*). The crude peptide was purified by $C_{18}$ reversed phase high-performance liquid chromatography (HPLC) and then analyzed by matrix-assisted laser desorption/ionization (MALDI) mass spectrometry. The molecular weight measured from mass spectrometry was 2140.39, which was consistent with the calculated value (2140.37).

*CD Spectroscopy.* Peptide solutions contained 0.50 mg/mL peptide in aqueous buffer containing 10 mM $Na_3PO_4$ (pH 6.2) and 0−50% trifluoroethanol (TFE). CD spectra were recorded at 15 °C in a 0.1 cm cuvette on a JASCO model 810 spectropolarimeter from 190 to 260 nm using a single scan. Spectra were baseline corrected by buffer subtraction and deconvoluted using the CDNN program to calculate the percentage of α-helix (*14*).

*NMR Spectroscopy.* Samples of the E6ap peptide were prepared at 2 mM in an aqueous buffer containing $Na_3PO_4$ and 0 or 40% deuterated TFE (Cambridge Isotope Laboratories, Andover, MA) at pH 6.2. Samples were locked on the deuterium signal from TFE or from 10% $D_2O$ that was added to aqueous samples. E6ap in $D_2O$ was prepared by lyophilizing the peptide solution and dissolving it in 99.96% $D_2O$ with TFE added to a final concentration of 40%.

Most NMR spectra were collected at 15 °C on a Bruker AMX-500 spectrometer with a proton frequency of 500.14 MHz. An additional NOESY spectrum was obtained on a home-built 750 MHz instrument at the Frances Bitter National Magnetic Laboratory, Massachusetts Institute of Technology (Cambridge, MA). The carrier frequency was set on the water resonance, which was suppressed using presaturation. Two-dimensional NOESY, total correlation spectroscopy (TOCSY), and double-quantum-filtered COSY spectra were collected with 2048 data points in $t_2$, 512 time-proportional phase increments in $t_1$, and a spectral width of approximately 7000 Hz (*15*). Spectra were processed with sine-bell window functions shifted by 45° in $t_2$ and 60° in $t_1$ using FELIX software. The NOESY spectrum was acquired with a mixing time of 100 ms. The TOCSY spectrum was recorded with a mixing time of 40 ms and an MLEV-17 mixing sequence (*15*). Final spectra were zero filled to 2048 × 1024 (real) points.

Proton assignments were made by following standard homonuclear methods (*16*). In brief, the amino acid spin systems were assigned first from the TOCSY and double-quantum COSY spectrum to establish intraresidue connectivities. Sequential assignment was then performed using the observed $d_{\alpha N}$, $d_{\beta N}$, and $d_{NN}$ NOE cross-peaks between sequential residues starting with the spin system of the single threonine residue.

NOESY cross-peak intensities were converted to distances and calibrated with published methods (*17, 18*). The three-bond coupling constants, $^3J$(HN,Hα), were measured from the splitting of cross-peaks in the HN dimension (*19*). Coupling constants were converted to dihedral restraints ($\phi$ angles) using the Karplus relationship. $\chi_1$ angle restraints were determined from $^3J$(Hα,H$\beta$), measured from splittings in the Hα dimension of spectra collected in a $D_2O$ solution, and by using H$\beta$−amide and H$\beta$−Hα NOE cross-peak intensities (*16*). There was no evidence of intermolecular NOE cross-peaks, and the CD spectra collected at two different concentrations (0.5 and 1.0 mg/mL) were essentially identical. Therefore, structure calculations were performed on a monomer of E6ap by employing 147 distance restraints and 17 dihedral restraints (10 $\phi$ and 7 $\chi_1$) that were entered into the XPLOR program (Biosym Technologies, San Diego, CA). A simulated annealing method was used with the PROLSQ force field to generate 30 structures, of which 23 were accepted (*20*, *21*). These structures superimposed with a root-mean-square (rms) deviation of 0.54 ± 0.17 Å for the backbone of residues 4−16, which had well-defined backbone torsion angles (*17*). Four of the excluded structures had grossly different structures (rms deviation of >0.9 Å), and three had scattered NOE violations of >0.2 Å. The largest NOE violation in the accepted structures was 0.15 Å. An average structure was created from the set of 23 and energy minimized using the PROLSQ force field (*21*). The hydrogen-bonding pattern was determined from the calculated structures using InsightII.

*Plasmids.* Plasmid pSPBPVE6 encodes BPV-1 E6 in a pSP65 vector. GST−E6AP and GST−E6APdl (deletion of amino acids 391−401) were described previously (*5*). Plasmid pSP16E6 was obtained from K. Vousden (*22*).

*Site-Directed Mutagenesis.* E6AP (76 kDa) cloned in the pGEX-2T vector served as the starting material for mutagenesis. E6AP mutants were made using the QuickChange Site-

Directed Mutagenesis Kit from Stratagene. The amino acid sequence of the E6-binding motif of the E6AP protein was renumbered according to the E6ap peptide numbering. The following primers, along with the corresponding complementary oligonucleotides, were used: L9A, 5′-CCAGCGAGCTGACAGCTCAGGAACTTTTGGG-3′; L12A, 5′-GACACTTCAGGAAGCTTTGGGAGAAG-3′; L13A, 5′-CTTCAGGAACTTGCGGGAGAAGAAAG-3′; E11P, 5′-GCTGACACTTCAGCCACTTTTGGGAG-3′; E11A, 5′-GCTGACACTTCAGGCACTTTTGGGAG-3′; G14A, 5′-GGAACTTTTGGCAGAAGAAAGAAG-3′; and L7K, 5′-GAGTCCAGCGAGAAGACACTTCAGG-3′.

*Protein Expression and Purification.* Glutathione *S*-transferase (GST) fusion proteins were expressed in *Escherichia coli* strain DH5α. One liter of LB medium was inoculated with 100 mL of stationary culture and the culture grown for 1 h before induction with 0.2 mM IPTG for 3 h. Cells were harvested by centrifugation, resuspended in 50 mL of low-salt association buffer (LSAB) [100 mM Tris-HCl (pH 8.0), 100 mM NaCl, 1% NP-40, and 1 mM phenylmethanesulfonyl fluoride] with 0.03% SDS and 2 mM DTT, and lysed by sonication. After centrifugation at 10000*g* for 10 min, the precipitate was resuspended in 8 M urea, combined with the supernatant, and mixed with glutathione−Sepharose beads (Pharmacia). The mixture was subjected to rotary shaking for 2 h at 4 °C. The beads were then collected by centrifugation at 1000*g* for 2 min, washed three times with 20 volumes of LSAB, and stored at 4 °C. In vitro-translated E6 proteins were prepared by using the rabbit reticulocyte lysate translation system (Promega) and $^{35}$S-labeled cysteine (ICN, Irvine, CA).

*In Vitro Association Experiment.* For in vitro binding, glutathione−Sepharose beads containing approximately 2 $\mu$g of GST−E6AP fusion proteins were combined with 1−5 $\mu$L of $^{35}$S-labeled in vitro-translated E6 proteins in LSAB in a total volume of 250 $\mu$L. The mixtures were gently shaken for 3 h at 4 °C. The mixtures were then washed extensively with LSAB, boiled in SDS-gel loading buffer, and electrophoresed on SDS−polyacrylamide gels. Gels were fixed and intensities determined by densitometry (Molecular Imager, Bio-Rad, Hercules, CA).

## RESULTS

*Spectral Analysis.* Primary sequence analysis of the peptide indicated a tendency for the peptide to form an α-helix. In aqueous solution, deconvolution of the CD spectrum indicated about 20% helix, indicating that three or four residues were in a helical conformation. A series of circular dichroism (CD) spectra collected with increasing percentages of TFE showed that 40% TFE provided most of the transition to helix (Figure 2A). The NMR spectra for the sample in aqueous solution exhibited a lack of resonance dispersion, indicating that the peptide was poorly structured. In 40% TFE, the two-dimensional NOESY spectrum exhibited the typical feature in helical peptides of strong NOE cross-peaks between sequential residues in the amide−amide proton region (Figure 3).

After the completion of resonance assignments, cross-peaks in the NOESY spectra were assigned. Short-range NOE interactions (between adjacent residues) and medium-range interactions (between residues separated by at least

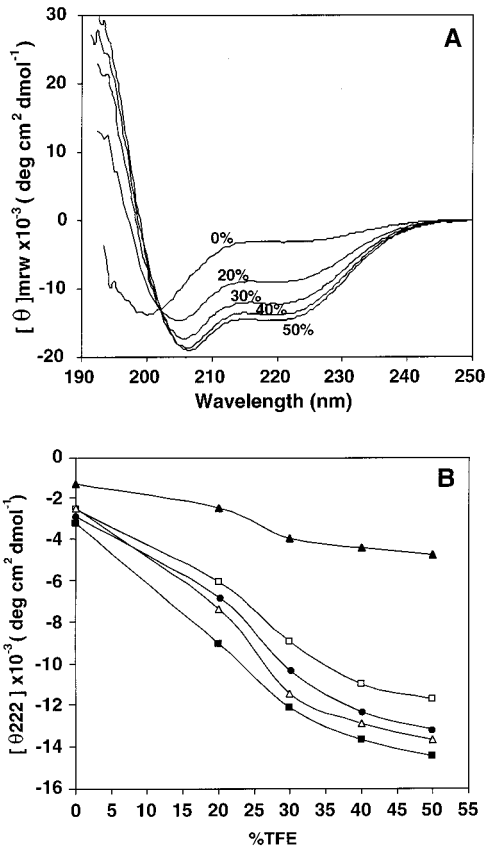


FIGURE 2: Circular dichroism spectroscopy of wild-type and mutant E6ap peptides. (A) CD spectra for wild-type E6ap are shown at 0, 20, 30, 40, and 50% TFE. (B) CD signal at 222 nm for wild-type and mutant peptides as a function of TFE concentration: wild-type E6ap (■), L9A (□), L13A (●), L12A (△), and E11P (▲).

two residues) were observed in the cross-peaks (Figure 4). Three-bond coupling constants between amide and Cα protons in the same residue were measured from the observed splitting of amide proton cross-peaks taken from one-dimensional slices from a resolution-enhanced two-dimensional NOESY spectrum (*23*). The $^3J$(HN,Hα) coupling constants for residues 5−12 were small (<6 Hz), which suggested that these residues had dihedral angles consistent with an α-helical conformation (Figure 4). NOE cross-peaks describing amide−amide proton contacts, α-amide proton contacts, and $\beta$-amide proton contacts also defined a helix for the peptide between residues 4 and 13 (Figure 4). Although the amide−amide NOE data and chemical shift data suggested an additional helical fragment between residues 13 and 17, the $\beta$-amide proton contacts and the $^3J$(HN,Hα) coupling constants indicated that the helical element of residues may be poorly formed. Due to the partial overlap in the spectrum, not all of the expected continuous NOE interactions between protons separated by three or more residues in the primary sequence could be measured with certainty. Besides those shown in Figure 4, we also observed an αH−NH ($i$,$i$+4) NOE peak between Leu 7 and Glu 11 and αH−$\beta$H ($i$,$i$+3) peaks between Ser 4 and Leu 7 and between Leu 7 and Gln 10. The observed chemical shifts of Cα protons for residues 4−13, which were shifted upfield relative to the standard chemical shifts observed for random coil peptides in 40% TFE, provide further evidence that the peptide was helical (*24*, *25*).

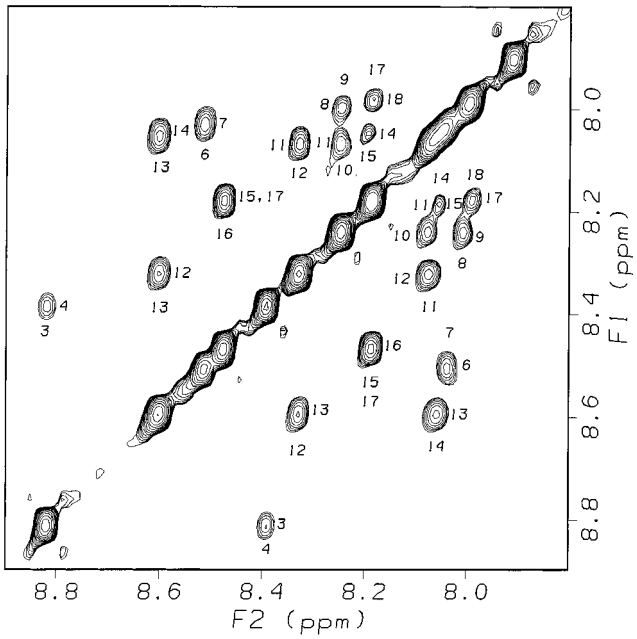


FIGURE 3: Amide−amide proton region of the two-dimensional NOESY spectrum of E6ap collected at 500 MHz. The peptide comprises amino acids 391−408 of the E6-binding protein, E6AP, which are renumbered 1−18 here. The cross-peaks between the amide proton of one amino acid and that of the amino acid next in sequence indicate an α-helical structure between residues 4 and 13. The sample comprised 2 mM peptide in 40% trifluoroethanol and 10 mM phosphate at pH 6.2 and 15 °C. The NOE mixing time was 100 ms.

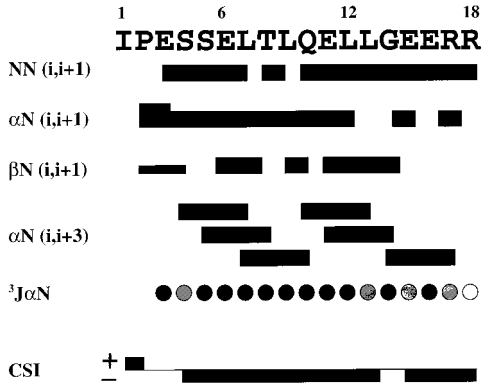


FIGURE 4: Summary of conformational data for the E6ap peptide (from analysis of 500 and 750 MHz NMR data). NOE peaks describing amide−amide (NN) proton contacts, α-amide (αN) proton contacts, *β*-amide (*β*N) proton contacts, and Hα(*i*)−H*β*-(*i*+3) contacts are illustrated, with the thickness of the filled-in bars corresponding to NOE intensity. Thin bars represent weak contacts (4.0−5.0 Å); medium bars represent medium contacts (3.0−4.0 Å), and thick bars represent strong contacts (2.0−3.0 Å). Coupling constants $^3J$(HN,Hα) are indicated for values of <6 Hz (●), of >7 Hz (○), and between 6 and 7 Hz (gray circles). The chemical shift index (CSI) is positive when the difference between the observed Hα chemical shift and that of a random coil peptide was greater than 0.1 ppm and negative when the difference was less than −0.1 ppm.

*Structure Determination of the E6ap Peptide.* NOESY cross-peak intensities were converted into distance restraints, and coupling constants were converted into dihedral restraints. A set of 147 distance restraints and 17 dihedral restraints were used to generate 23 final structures (Figure 5A). These structures had negative total energies of −227 ± 26 kcal/mol.

Overall, the calculated structures of the E6ap peptide were consistent with the NMR data. The three residues at the N-terminus and the two residues at the C-terminus were not well ordered. Overall, the apparent structure is one helix starting near residue 4 and continuing to the C-terminus (Figure 5B). Closer inspection revealed two families of structures. Seven of the 23 structures (family 1) exhibited a continuous α-helix from residue 3 to the C-terminal end, as defined by the $\phi$ (approximately −60°) and $\psi$ (approximately −40°) angles. The majority of structures (family 2) exhibited an α-helix between residues 4 and 13 and a looser helix from residue 13 to the C-terminus, consistent with the glycine residue acting as the breaking (or distortion) point for the helical structure. For family 1 (seven members), the rms deviation for the backbone atoms to the average structure of family 1 was 0.4 ± 0.15 Å. For family 2 (16 members), the rms deviation for the backbone atoms to the average structure of family 2 was also 0.4 ± 0.15 Å. Because we could not distinguish between the two families from the NMR data, an average was taken over all the calculated structures. The energy-minimized average structure had a slightly higher total energy (−171 kcal/mol). All 23 of the calculated structures were superimposed on the average structure using residues with well-defined $\phi$ and $\psi$ torsion angles (Figure 5A). The rms deviation for residues 4−16 was 0.54 ± 0.17 Å for backbone atoms and 1.19 ± 0.23 Å for all heavy atoms.

Three leucines (Leu 9, Leu 12, and Leu 13) of the E6ap peptide were on the same surface of the helix (Figure 5). Sequence homology analysis shows that these residues are conserved among several E6-binding proteins. A search using a peptide library for E6-binding sequences also found sequences with hydrophobic residues at the positions corresponding to these leucines (*8*). From prior mutational analysis on a homologous protein (*12*) and our mutational results described below, these hydrophobic residues are likely to provide the most important interactions for the binding of E6AP to the E6 protein.

In the structure of E6ap, Gln 10 and Glu 15 are not on the same face as the leucines and Glu 11 is found on the opposite face of the helix. Gln 10 and Glu 15 are conserved as amino acids capable of hydrogen bonding (Gln, Glu, Asp, or Asn) among the E6-binding proteins (Figure 1). Interestingly, these residues were not determined in the peptide library assay to be important for binding to E6 (*8*). Glu 11 is conserved in both E6-binding proteins and the peptide library. The conservation and mutational analysis (see below) suggest a preference, but not a requirement, for hydrogen-bonding amino acid residues at these positions for the ability of the E6-binding proteins to interact with E6 (*8*). The structure of the E6ap peptide while bound to E6 would give a more precise definition of the roles played by each amino acid residue in contacting E6. In addition, the complex would show the bound conformation of the E6ap peptide. However, recombinant HPV-16 and HPV-18 E6 is poorly soluble and unstable (data not shown), and the limited dispersion of the E6ap peptide in aqueous solution would indicate that a transferred NOE experiment would be difficult to perform (*26*).

*Site-Directed Mutagenesis of the E6ap Peptide.* A mutational analysis was performed to determine the importance of individual amino acids of the peptide for E6 binding. On the basis of the solution structure of E6ap, single-amino acid

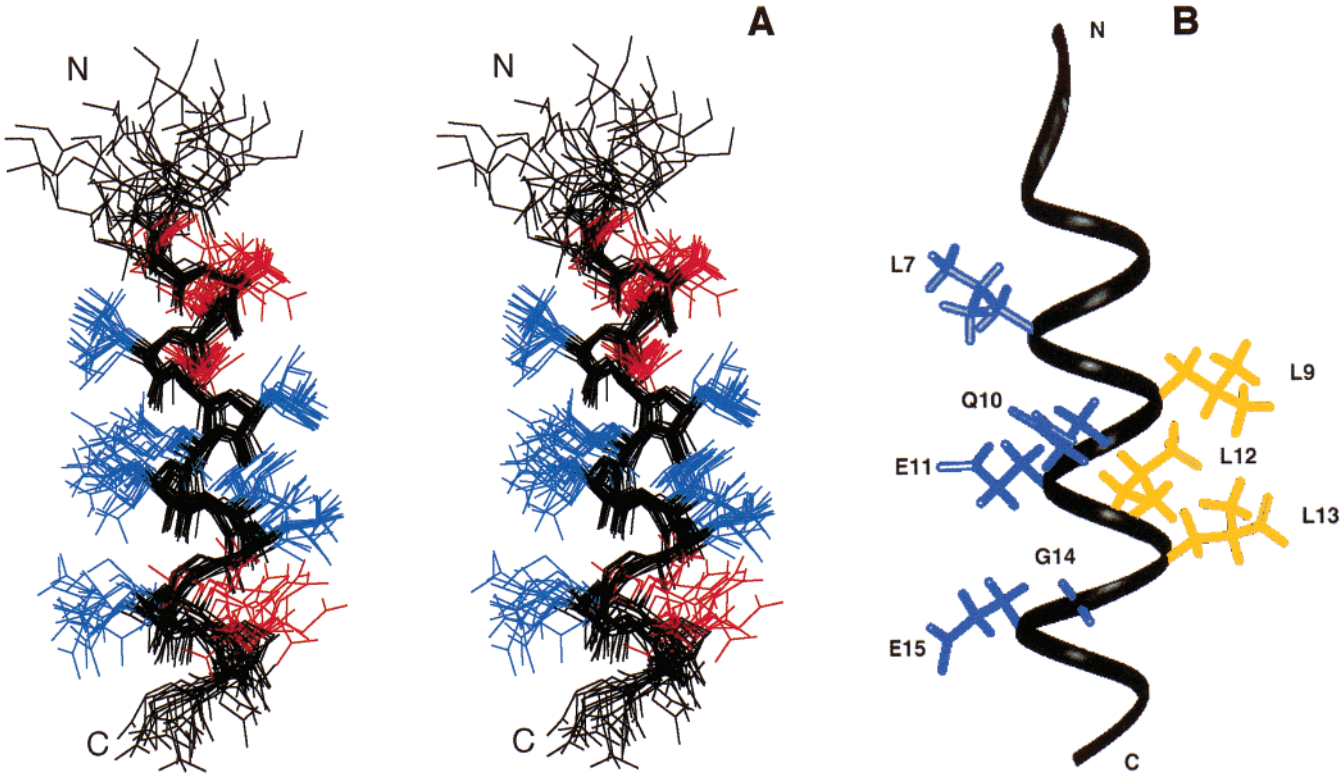


FIGURE 5: Solution structure of the E6ap peptide. (A) Stereoview overlay of the calculated structures. The backbone atoms for 23 structures were superimposed for residues 4−16. The root-mean-square (rms) deviation for the backbone atoms is 0.54 ± 0.17 Å. Side chains are shown only for well-defined residues. Side chains of the residues probed by mutagenesis are in blue; other side chains are shown in red. For clarity, only heavy atoms are shown. (B) Ribbon diagram of the average structure highlighting essential structural features. Hydrophobic residues (Leu 9, Leu 12, and Leu 13) that are essential for E6 binding are in gold. Other amino acid residues that were probed by mutagenesis are in blue.

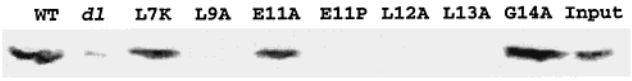


FIGURE 6: Conserved α-helical motif that is important for association with E6. In a representative experiment, glutathione−Sepharose beads containing GST−E6AP fusion proteins were combined with $^{35}$S-labeled in vitro-translated BPV-1 E6 proteins. After incubation and washes, the bound products were separated by SDS−polyacrylamide gel electrophoresis. Bands were realigned for qualitative comparison and were quantified by densitometry (Table 1). Input, loaded directly into the well, represents 5% of the $^{35}$S-labeled E6 used in each binding reaction.

mutations (L7K, L9A, E11P, E11A, L12A, L13A, and G14A) were chosen and constructed in the context of the nearly full-length 76 kDa E6AP protein. The mutant E6AP proteins were tested for their ability to bind BPV E6 using an in vitro association assay.

The peptide structure indicated that Leu 9, Leu 12, and Leu 13 were important for E6 binding. Substitution of any of these leucines with alanine completely abolished E6 binding (Figure 6 and Table 1). The replacement of a residue outside the consensus binding region (Leu 7) with Lys did not significantly affect binding to E6. From the structure, Glu 11 does not appear to be involved in E6 binding. Consistent with this, the Glu 11 to Ala mutation did not affect the binding to E6, although the replacement of Glu 11 with proline abolished E6 binding. The Gly 14 to Ala mutation had little effect on E6 binding (Figure 6). E6AP was also assayed for binding to HPV-16 E6. Similar effects for all mutations were observed for both the bovine and human types of virus (Table 1).

Table 1: Construction of Point Mutations for GST−E6AP Fusion Proteins

| Name | Sequence[a] | BPV-1[b] % binding | HPV-16[b] % binding |
|---|---|---|---|
| Wild-type | IPESS ELTLQ ELLGE ERR | (100) | (100) |
| Consensus | ----- ---Lh hϕLGh --- | | |
| L7K | IPESS E**K**TLQ ELLGE ERR | 47±16 | 50±14 |
| L9A | IPESS ELT**A**Q ELLGE ERR | 0±0 | 0±0 |
| E11A | IPESS ELTLQ **A**LLGE ERR | 62±21 | 50±17 |
| E11P | IPESS ELTLQ **P**LLGE ERR | 0±0 | 0±0 |
| L12A | IPESS ELTLQ E**A**LGE ERR | 1±2 | 0±0 |
| L13A | IPESS ELTLQ EL**A**GE ERR | 0±0 | 0±0 |
| G14A | IPESS ELTLQ ELL**A**E ERR | 115±53 | 80±6 |

[a] The sequences of the 18 amino acids (amino acids 391−408) in the E6 binding domain of E6AP are given. The consensus motif comprises seven amino acid residues (h is D, E, Q, or N; $\phi$ is hydrophobic) that are conserved for at least four of the five E6-binding proteins (see Figure 1). [b] The percent binding relative to wild-type E6AP is shown as the average of two independent determinations.

We examined the structures of the mutants that had a signficant loss in the extent of binding of E6 using CD spectroscopy (Figure 2B). In the absence of TFE, wild type E6ap and the mutants L9A, L21A, and L13A had almost the same helix content (15−20%). They all underwent a

similar transition between 20 and 30% TFE, and all exhibited a helix content between 40 and 50% in 50% TFE. Therefore, the structures of the mutant peptides are very similar to that of the wild type. In contrast, the E11P peptide was significantly different, both at 0% TFE with a low signal intensity that deconvoluted to about 5% helix and at 50% TFE where the helix content was only 25%.

## DISCUSSION

All HPV-positive cervical carcinomas express E6 protein from high-risk human papillomaviruses (*3*). E6 is an oncoprotein without known intrinsic enzymatic activity. Instead, its biological effects are mediated by interactions with human cellular factors such as E6AP, E6BP, tuberin, paxillin, and IRF-3 (*3*). These proteins share a homologous E6-binding sequence comprising seven amino acid residues (Figure 1).

In this study, we determined the structure of a peptide representing the E6 binding domain (E6ap) of E6AP. We observed that the well-defined helix in the E6ap peptide began at residue 4 and ended at residue 13. The helix included the most conserved part of the E6-binding consensus region, consistent with the low-resolution structure determined for E6BP structure and site-directed mutagenesis (*12*). Previous studies have shown that a single peptide inhibits the interaction between E6 and both E6AP and E6BP (*8*). This suggests that both proteins interact with E6 with a similar domain.

Typical of most small peptides, E6ap was only partly helical in aqueous solution, but was observed to form a stable α-helix in the presence of 40% trifluoroethanol (TFE). Although TFE can stabilize the helix for a peptide with an intrinsic helical structure, the reagent is assumed not to induce one artificially (*13*). Typically, short peptides with intrinsic helical propensities are unstructured in aqueous solution or in the absence of a partner protein (*27*), but form α-helices in the presence of TFE. TFE is thought to mimic the protein partner in the sense that it stabilizes helices in regions with intrinsic α-helical propensity that are likely to form helices when binding to their protein partner. For example, a helical structure was observed for the synthetic peptide encompassing the calmodulin-binding domain of skeletal muscle myosin light kinase both in the presence of TFE and in the complex with calmodulin (*28*). Because there are examples where TFE appears to artificially induce a helix (*29*, *30*), the structure was validated by an independent method such as site-directed mutagenesis.

Mutations were chosen to test the E6ap peptide structure and also were designed to answer two questions. (1) Are the hydrophobic residues on the same side of the helix important for E6 binding? (2) Is the helical structure necessary for E6 binding? Site-directed mutagenesis confirmed the requirements of the conserved hydrophobic amino acids for binding to E6 (Figure 6), and CD spectroscopy showed that the structures of these mutants were not significantly perturbed (Figure 2B). Mutation of Glu 11 to alanine showed little change in binding to E6, suggesting that the side chain is not necessary for the binding of E6AP to the E6 protein. On the other hand, the E11P mutation abolished binding. Consistent with proline generally breaking or distorting an α helix, its introduction significantly destabilized the E6ap peptide (Figure 2B). Breaking of the helix at residue 11 is likely to misalign the residues important for binding to E6. The result also suggests that the E6-binding motif of the E6AP protein must be helical while bound to E6.

The E6 protein of BPV-1 and HPV-16 E6 are basic proteins with estimated isoelectric points of 8.7 and 9.0, respectively. Therefore, it might be expected that the charge on the E6 interacting peptide would influence binding. However, the mutation of Leu 7 to lysine, which is outside the consensus binding motif, did not significantly affect the binding to E6. Although the aliphatic portion of the lysine residue may substitute, in part, for the hydrophobic leucine side chain, it is also likely that a hydrophobic residue at this position is not required (Figure 6).

Gly 14 is conserved in E6-binding proteins and is observed in a library of peptides selected for interaction with E6. In paxillin, alanine takes the place of glycine in the corresponding position. In E6AP, the Gly 14 to alanine mutation did not significantly affect binding to E6. The conservation of an amino acid with a small side chain (glycine or alanine) suggests that it is a required feature of proteins that bind E6. The small amino acid residue may allow for conformational changes in the peptide upon binding to E6, or may be required sterically for the binding motif to fit into its binding pocket on E6.

A model for the binding pocket of E6 can be derived from the structure of E6ap and the mutagenesis results. This pocket is imagined to be deep, with hydrophobic amino acid residues lining the bottom that interact with Leu 9, Leu 12, and Leu 13 of E6ap. The presence of a small amino acid residue at position 14 indicates a steric requirement that limits the sequences that can bind E6.

The structure and mutagenesis studies of the E6ap peptide and E6bp peptides and the homology analysis of several E6-binding proteins suggest a general E6-binding motif. The reproducibility of the mutagenesis results for binding to both BPV-1 and HPV-16 E6 proteins indicates not only that the E6-binding motif is general in the sense that it is present on several proteins that bind E6 but also that the E6-binding motif is sufficiently general to interact with the E6 from any papillomaviral type. This E6-binding motif is an α-helix, and the residues used for contacting E6 are on the same surface of the helix forming a hydrophobic patch, with contributing roles played by hydrogen bond-forming amino acid residues and steric requirements surrounding the small amino acid residue.

The E6-binding motif falls into the LXXLL motif family, a general protein−protein interaction motif for which the contact is made through a hydrophobic patch formed by the leucine residues that packs into a hydrophobic pocket of its receptor (*31*−*33*). The three-dimensional structures of LXXLL motifs show them to be helical while bound to the ligand-binding domains of their receptors but mostly unstructured when unbound (*34*−*36*). The folding transition of the LXXLL motif may contribute to the specificity of molecular recognition (*27*, *37*). In eukaryotic cells, protein−protein interactions that are involved in cellular regulation are often mediated through such short helical motifs (*38*). Thus, the viral E6 protein binds human cellular proteins by using a mechanism similar to that used by cellular proteins in interacting with each other.

## ACKNOWLEDGMENT

We thank Dr. Jim Sudmeier for assistance with setting up the NMR experiments.

## SUPPORTING INFORMATION AVAILABLE

Table of chemical shift assignments. This material is available free of charge via the Internet at http://pubs.acs.org.

## REFERENCES

1. Scheffner, M., Romanczuk, H., Munger, K., Huibregtse, J. M., Mietz, J. A., and Howley, P. M. (1994) *Curr. Top. Microbiol. Immunol. 186*, 83−99.
2. Alani, R. M., and Munger, K. (1998) *J. Clin. Oncol. 16*, 330−337.
3. Rapp, L., and Chen, J. J. (1998) *Biochim. Biophys. Acta 1378*, F1−F19.
4. Scheffner, M., Huibregtse, J. M., Vierstra, R. D., and Howley, P. M. (1993) *Cell 75*, 495−505.
5. Huibregtse, J. M., Scheffner, M., and Howley, P. M. (1993) *Mol. Cell. Biol. 13*, 4918−4927.
6. Malzac, P., Webber, H., Moncla, A., Graham, J. M., Kukolich, M., Williams, C., Pagon, R. A., Ramsdell, L. A., Kishino, T., and Wagstaff, J. (1998) *Am. J. Hum. Genet. 62*, 1353−1360.
7. Chen, J. J., Reid, C. E., Band, V., and Androphy, E. J. (1995) *Science 269*, 529−531.
8. Elston, R. C., Napthine, S., and Doorbar, J. (1998) *J. Gen. Virol. 79*, 371−374.
9. Vande Pol, S. B., Brown, M. C., and Turner, C. E. (1998) *Oncogene 16*, 43−52.
10. Brown, M. C., Curtis, M. S., and Turner, C. E. (1998) *Nat. Struct. Biol. 5*, 677−678.
11. Ronco, L. V., Karpova, A. Y., Vidal, M., and Howley, P. M. (1998) *Genes Dev. 12*, 2061−2072.
12. Chen, J. J., Hong, Y., Rustamzadeh, E., Baleja, J. D., and Androphy, E. J. (1998) *J. Biol. Chem. 273*, 13537−13544.
13. Roy, R. S., Kim, S., Baleja, J. D., and Walsh, C. T. (1998) *Chem. Biol. 5*, 217−228.
14. Bohm, G., Muhr, R., and Jaenicke, R. (1992) *Protein Eng. 5*, 191−195.
15. Cavanagh, J., Fairbrother, W. J., Palmer, A. G., III, and Skelton, N. J. (1996) *Protein NMR spectroscopy*, Academic Press, San Diego.
16. Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
17. Hyberts, S. G., Goldberg, M. S., Havel, T. F., and Wagner, G. (1992) *Protein Sci. 1*, 736−751.
18. Detlefsen, D. J., Thanabal, V., Pecoraro, V. L., and Wagner, G. (1991) *Biochemistry 30*, 9040−9046.
19. Wang, Y., Nip, A. M., and Wishart, D. S. (1997) *J. Biomol. NMR 10*, 373−382.
20. Caldwell, J. E. (1998) Ph.D. Thesis, University of Wisconsin, Madison, WI.
21. Linge, J. P., and Nilges, M. (1999) *J. Biomol. NMR 13*, 51−59.
22. Crook, T., Tidy, J. A., and Vousden, K. H. (1991) *Cell 67*, 547−556.
23. Szyperski, T., Guntert, P., Otting, G., and Wuthrich, K. (1992) *J. Magn. Reson. 99*, 552−560.
24. Merutka, G., Dyson, H. J., and Wright, P. E. (1995) *J. Biomol. NMR 5*, 14−24.
25. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) *J. Mol. Biol. 222*, 311−333.
26. Ni, F., and Scheraga, H. A. (1994) *Acc. Chem. Res. 27*, 257−264.
27. Wright, P. E., and Dyson, H. J. (1999) *J. Mol. Biol. 293*, 321−331.
28. Zhang, M., Yuan, T., and Vogel, H. J. (1993) *Protein Sci. 2*, 1931−1937.
29. Chaloin, L., Vidal, P., Heitz, A., Van Mau, N., Méry, J., Divita, G., and Heitz, F. (1997) *Biochemistry 36*, 11179−11187.
30. Najbar, L. V., Craik, D. J., Wade, J. D., Salvatore, D., and McLeish, M. J. (1997) *Biochemistry 36*, 11525−11533.
31. Heery, D. M., Kalkhoven, E., Hoare, S., and Parker, M. G. (1997) *Nature 387*, 733−736.
32. Torchia, J., Rose, D. W., Inostroza, J., Kamei, Y., Westin, S., Glass, C. K., and Rosenfeld, M. G. (1997) *Nature 387*, 677−684.
33. McInerney, E. M., Rose, D. W., Flynn, S. E., Westin, S., Mullen, T. M., Krones, A., Inostroza, J., Torchia, J., Nolte, R. T., Assa-Munt, N., Milburn, M. V., Glass, C. K., and Rosenfeld, M. G. (1998) *Genes Dev. 12*, 3357−3368.
34. Darimont, B. D., Wagner, R. L., Apriletti, J. W., Stallcup, M. R., Kushner, P. J., Baxter, J. D., Fletterick, R. J., and Yamamoto, K. R. (1998) *Genes Dev. 12*, 3343−3356.
35. Nolte, R. T., Wisely, G. B., Westin, S., Cobb, J. E., Lambert, M. H., Kurokawa, R., Rosenfeld, M. G., Willson, T. M., Glass, C. K., and Milburn, M. V. (1998) *Nature 395*, 137−143.
36. Shiau, A. K., Barstad, D., Loria, P. M., Cheng, L., Kushner, P. J., Agard, D. A., and Greene, G. L. (1998) *Cell 95*, 927−937.
37. Spolar, R. S., and Record, M. T., Jr. (1994) *Science 263*, 777−784.
38. Uesugi, M., and Verdine, G. L. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 14801−14806.

BI0019592